

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 24, 2003, 15:43:18 : Search time 63.5263 Seconds
(without alignments)
2848.398 Million cell updates/sec

Title: US-09-974-973a-19

Perfect score: 5788

Sequence: 1 MSTHTSSTLPFRKILVANR.....RVVYPATKVGSDLIWVS 1140

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A.GeneSeq_19Jun03.*

```

1: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	5788	100.0	1140	22	AAG90511
2	5788	100.0	1140	22	AAB67129
3	5788	100.0	1140	23	AAE25601
4	5788	100.0	1140	23	AAU98053
5	5784	99.9	1141	21	AAB01436
6	5782	99.9	1140	20	AAW35971
7	5780	99.9	1140	22	AAU93249
8	5759	99.5	1140	23	AAU98052
9	5759	99.5	1157	23	AAU98050

10	5302.5	91.6	1139	22	AAB83180	Corynebacterium th
11	2621	45.3	532	22	AAB79302	Corynebacterium gl
12	2524	43.6	1148	22	AAU00511	Listellus subtilis
13	2493	43.1	1146	23	ABB47612	Staphylococcus aur
14	2488	43.0	1147	22	AAU33972	Enterococcus faeca
15	2472.5	42.7	1142	22	AAU35213	S. epidermidis ope
16	2464	42.6	1151	22	AAU38583	Staphylococcus epi
17	2464	42.6	1154	22	ABB38583	Drosophila melanog
18	2441	42.2	1181	22	ABB58211	Drosophila melanog
19	2441	42.2	1181	22	ABB66604	Drosophila melanog
20	2441	42.2	1181	22	ABB67309	Drosophila melanog
21	2433.5	42.0	1196	22	ABB67309	Lactococcus lactis
22	2405	41.6	1136	23	ABB53980	Corynebacterium gl
23	2357	40.7	461	22	AAB79303	Corynebacterium gl
24	2299	39.7	1073	22	AAU36768	Staphylococcus aur
25	1546.5	26.7	320	22	AAB79300	Corynebacterium gl
26	1546.5	26.7	320	22	AAB79301	Corynebacterium gl
27	1375	23.8	272	22	AAB79298	Corynebacterium gl
28	1375	23.8	272	22	AAB79299	Corynebacterium gl
29	1036.5	17.9	447	17	AAW05207	Anabaena biotin bi
30	1036.5	17.9	447	19	AAW70402	Anabaena biotin ca
31	1036.5	17.9	448	15	AAR51080	Biotin carboxylase
32	1004.5	17.4	453	15	AAR51083	fabg gene encoding
33	1004.5	17.4	453	17	AAW05208	Synechococcus biot
34	1004.5	17.4	453	19	AAW70403	Synechococcus biot
35	978	16.9	425	21	AAB11663	A. vitis hypersens
36	976.5	16.9	471	22	AAU33719	Pseudomonas aerugi
37	957	16.5	449	22	AAU38292	Salmonella typhi c
38	953.5	16.5	456	23	ABP28017	Lactococcus poly
39	951.5	16.4	455	23	ABB54091	Lactococcus lactis
40	946	16.3	448	22	AAU35567	Haemophilus influe
41	942.5	16.3	455	22	AAU38024	Streptococcus pneu
42	942.5	16.3	455	24	ABU00779	S. pneumoniae type
43	941.5	16.3	455	22	AAU37771	Streptococcus pneu
44	939.5	16.2	455	22	AAW01054	CPE 57 protein seq
45	938	16.2	449	16	AAR66742	Biotin-carboxylase

ALIGNMENTS

```

RESULT 1
AAG90511
ID: AAG90511 strand: Protein; 1140 AA.
XX
AC AAG90511;
XX
26-SEP-2001 (first entry)
XX
DE C glutamicum protein fragment SEQ ID NO: 4265.
XX
KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
XX organic acid synthesis.
XX
OS Corynebacterium glutamicum.
XX
PN EPI108790-A2.
XX
PD 20-JUN-2001.
XX
PF 18-DEC-2000; 2000EP-0127688.
XX
PR 16-DEC-1999; 99AP-0377484.
XX 07-APR-2000; 2000UP-0159162.
XX 03-AUG-2000; 2000UP-0280988.
XX
PA (KYOW ) KYOWA HAKKO KOGYO KK.
XX
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
XX Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX WPI: 2001-376931/40.
XX
DR N-PSDB; AAB65730.

```


xx	Sequence	1140 AA:	100.0%: Score 5788; DB 22; Length 1140;
sq	Match	Best Local Similarity 100.0%: Pred. No. 0;	Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MSTHTSSTLPFAKKILVANRGEIAVRAFRALLETGAATVAIYPREDRGSFHRFSASEAVR	100.0%: Score 5788; DB 22; Length 1140;	60
DB	1 MSTHTSSTLPFAKKILVANRGEIAVRAFRALLETGAATVAIYPREDRGSFHRFSASEAVR	100.0%: Score 5788; DB 22; Length 1140;	60
QY	61 IGTGGSPPKALVDDEITIGAAKKKKADAIYGYGFLSNAQLARCAENGITFTGPTPEV	100.0%: Score 5788; DB 22; Length 1140;	120
DB	61 IGTGGSPPKALVDDEITIGAAKKKKADAIYGYGFLSNAQLARCAENGITFTGPTPEV	100.0%: Score 5788; DB 22; Length 1140;	120
QY	121 LDLTGDKSRATTAARKKAGLPVLAISTPSKNIIDEIYKSAEGOTYPTFVAAVVGGGGRGRNF	100.0%: Score 5788; DB 22; Length 1140;	180
DB	121 LDLTGDKSRATTAARKKAGLPVLAISTPSKNIIDEIYKSAEGOTYPTFVAAVVGGGGRGRNF	100.0%: Score 5788; DB 22; Length 1140;	180
QY	181 VASPEDELKLTAEASREAEAFGCGAVYVERAVINPOHIEVOILGDHTGEVVIHYERDCS	100.0%: Score 5788; DB 22; Length 1140;	240
DB	181 VASPEDELKLTAEASREAEAFGCGAVYVERAVINPOHIEVOILGDHTGEVVIHYERDCS	100.0%: Score 5788; DB 22; Length 1140;	240
QY	241 LQRRHQKVELTAPQOHDPBELDRICADAVKFCNSIGYQAGYVELYDEKGNHYFTMN	100.0%: Score 5788; DB 22; Length 1140;	300
DB	241 LQRRHQKVELTAPQOHDPBELDRICADAVKFCNSIGYQAGYVELYDEKGNHYFTMN	100.0%: Score 5788; DB 22; Length 1140;	300
QY	301 PRIQEVHTVEEVT EVDLVKQMRILAAGATLKEIGLTPDDKIKTHGAALQCRITTPDPNNG	100.0%: Score 5788; DB 22; Length 1140;	360
DB	301 PRIQEVHTVEEVT EVDLVKQMRILAAGATLKEIGLTPDDKIKTHGAALQCRITTPDPNNG	100.0%: Score 5788; DB 22; Length 1140;	360
QY	361 FRPDTGITTAARSPGAGVRLDGAALQGEITTAHEDSMVLVMTKCGSFETAVARAKORAL	100.0%: Score 5788; DB 22; Length 1140;	420
DB	361 FRPDTGITTAARSPGAGVRLDGAALQGEITTAHEDSMVLVMTKCGSFETAVARAKORAL	100.0%: Score 5788; DB 22; Length 1140;	420
QY	421 AEFYVSGVATNIGFRLALREEDTSKRIANGFIADHPLLQAPPADDEQGRILDYLDVY	100.0%: Score 5788; DB 22; Length 1140;	480
DB	421 AEFYVSGVATNIGFRLALREEDTSKRIANGFIADHPLLQAPPADDEQGRILDYLDVY	100.0%: Score 5788; DB 22; Length 1140;	480
QY	481 TVNNPHGRPPDVAAPIDKLPINIDPLPRGSRBDLKOLGPAAFRDLREDDALAVPTT	100.0%: Score 5788; DB 22; Length 1140;	540
DB	481 TVNNPHGRPPDVAAPIDKLPINIDPLPRGSRBDLKOLGPAAFRDLREDDALAVPTT	100.0%: Score 5788; DB 22; Length 1140;	540
QY	541 FRDAHQSLLATRVBSFALKPAAEAVAKLTPELLSYEAMGAGTYDVAARFLEDDPMDRLDE	100.0%: Score 5788; DB 22; Length 1140;	600
DB	541 FRDAHQSLLATRVBSFALKPAAEAVAKLTPELLSYEAMGAGTYDVAARFLEDDPMDRLDE	100.0%: Score 5788; DB 22; Length 1140;	600
QY	601 LREAMPNVNIMOLRGRTVGYTPYPDSCVCAFYKEAASGCVDFRITDALNDVSOQMPRA	100.0%: Score 5788; DB 22; Length 1140;	660
DB	601 LREAMPNVNIMOLRGRTVGYTPYPDSCVCAFYKEAASGCVDFRITDALNDVSOQMPRA	100.0%: Score 5788; DB 22; Length 1140;	660
QY	661 IDAVLENTTAAAEVMAAYSGDLSDPNEKLYTLDDYLKMAEFTVSGAHTILAKDWAGLLR	100.0%: Score 5788; DB 22; Length 1140;	720
DB	661 IDAVLENTTAAAEVMAAYSGDLSDPNEKLYTLDDYLKMAEFTVSGAHTILAKDWAGLLR	100.0%: Score 5788; DB 22; Length 1140;	720
QY	721 PAAVTKLVTALARREEDLPVHVHTHDTAGGOLATYFAAAQAGADAVDASAPLSGTTSOPS	100.0%: Score 5788; DB 22; Length 1140;	780
DB	721 PAAVTKLVTALARREEDLPVHVHTHDTAGGOLATYFAAAQAGADAVDASAPLSGTTSOPS	100.0%: Score 5788; DB 22; Length 1140;	780
QY	781 LSAIYAAFAHTRRDTGUSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEITPGOL	100.0%: Score 5788; DB 22; Length 1140;	840
DB	781 LSAIYAAFAHTRRDTGUSLEAVSDLEPYWEAVRGLYLPFESGTPGPTGRVYRHEITPGOL	100.0%: Score 5788; DB 22; Length 1140;	840
QY	841 SNLRAQATALGLADDFELIEDNYAAVNBMGLRPKTVTPSSKYVGDLAHLVGAGVDPADF	100.0%: Score 5788; DB 22; Length 1140;	900
DB	841 SNLRAQATALGLADDFELIEDNYAAVNBMGLRPKTVTPSSKYVGDLAHLVGAGVDPADF	100.0%: Score 5788; DB 22; Length 1140;	900
QY	901 AADPOKVDIPDSVIAFLGELGELNPPGGRPEPLRTALEGRSEGGAKPLTEVEPEEDAHDA	100.0%: Score 5788; DB 22; Length 1140;	960
DB	901 AADPOKVDIPDSVIAFLGELGELNPPGGRPEPLRTALEGRSEGGAKPLTEVEPEEDAHDA	100.0%: Score 5788; DB 22; Length 1140;	960
QY	961 DQSKERRNSLRLLPKPTEEFLEHRRRFGNTSALDRFEFYGLVEGRETLIRLPDVATP	100.0%: Score 5788; DB 22; Length 1140;	1020
DB	961 DQSKERRNSLRLLPKPTEEFLEHRRRFGNTSALDRFEFYGLVEGRETLIRLPDVATP	100.0%: Score 5788; DB 22; Length 1140;	1020

DB	961	DDSKERRSLNKLRLFPKPTTEFLHRRRRFGTSLDDREFFYGLVBEGRFLIRLPVKT	1020
QY	1021	LLVRDAISEPDDKGMRRVYVANGQIRPMKVRDSVESYATAEKADSSNKGHVAAPFA	1080
DB	1021	LLVRDAISEPDDKGMRRVYVANGQIRPMKVRDSVESYATAEKADSSNKGHVAAPFA	1080
QY	1081	GVATVTVAEGDEVKAGDAVAITLTKMKMATITTSVDGKIDRVVYPAATKYEGGDLIVYS	1140
DB	1081	GVATVTVAEGDEVKAGDAVAITLTKMKMATITTSVDGKIDRVVYPAATKYEGGDLIVYS	1140
RESULT 3			
ID	AAE25601	AAE25601 standard; Protein; 1140 AA.	
XX	AAE25601;		
AC	AAE25601;		
XX	04-NOV-2002	(first entry)	
DT	04-NOV-2002	(first entry)	
XX	Corynebacterium glutamicum pyruvate carboxylase protein.		
DE	Corynebacterium glutamicum pyruvate carboxylase protein.		
XX	Pyruvate carboxylase; anaplerotic enzyme; industrial fermentation;		
KW	oxaloacetate; growth; enzyme.		
XX	Corynebacterium glutamicum.		
OS	Corynebacterium glutamicum.		
XX	Key	Location/Qualifiers	
FX	Misc-difference 1	/note= "Encoded by GTG"	
FT	US6403351-B1.		
PN	11-JUN-2002.		
XX	03-OCT-2000; 2000US-0677575.		
PE	23-DEC-1998; 98US-0220081.		
XX	(ARCH) ARCHER-DANIELS MIDLAND CO.		
XX	Sinskey AJ, Lessard PA, Willis LB;		
PI	WPI: 2002-536037/57.		
XX	N-PSDB: AAD42046.		
DR	Novel pyruvate carboxylase polypeptide, useful for replenishing		
XX	oxaloacetate consumed for biosynthesis during growth, or lysine and		
XX	glutamic acid production in industrial fermentation -		
PT	Claim 1; Column 29-36; 28pp; English.		
PT	The present invention relates to novel pyruvate carboxylase proteins and		
XX	polynucleotides encoding such proteins. Sequences of the invention are		
CC	important anaplerotic enzymes for replenishing oxaloacetate consumed for		
CC	biosynthesis during growth, or lysine and glutamic acid production in		
CC	industrial fermentation. The present sequence is C. glutamicum pyruvate		
CC	carboxylase protein.		
XX	Sequence 1140 AA;		
XX	Query Match	100.0%; Score 5788; DB 23; Length 1140;	
XX	Best Local Similarity	100.0%; Pred. No. 0;	
XX	Matches 1140; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	MSTRHSSTLPFAFKILVANRGEIIVRAFRALLEGAAATVAIYPREDGSPHRSASAVR	60
DB	1	MSTRHSSTLPFAFKILVANRGEIIVRAFRALLEGAAATVAIYPREDGSPHRSASAVR	60
QY	61	IGTGESPKVALIDIDEITIGAAKKVKAADIYGYGFLESNQALRECAENGITFGPTPEV	120
DB	61	IGTGESPKVALIDIDEITIGAAKKVKAADIYGYGFLESNQALRECAENGITFGPTPEV	120
QY	121	LDLGGDSRAVTAAKKAGLPVLAESTSKNIDELIVKSAEGQYPIPVKAVAGGGGRMR	180

```

Db      121 LDLTGKSRSAVTAAKAGLPVLAESTPSKNIDEIVKSAEGQYPIEVKAAVAGGGRMFR 180
QY      181 VASPELRLKLTASREAEAFGDAVYVERAVINPOHIEVOILGHTGEVHLYERDCS 240
Db      181 VASPELRLKLTASREAEAFGDAVYVERAVINPOHIEVOILGHTGEVHLYERDCS 240
QY      241 LORRHOKVVEIAPAOHLDPDLRDRICADAVKFCRSIGYGAGTVEFLVDEKGNHVFLEM 300
Db      241 LORRHOKVVEIAPAOHLDPDLRDRICADAVKFCRSIGYGAGTVEFLVDEKGNHVFLEM 300
QY      301 PRIQVHTYEEETEVDLVKAQMRLAGATLKLGLTQDKIKTHGAALOCRTTEDPNNG 360
Db      301 PRIQVHTYEEETEVDLVKAQMRLAGATLKLGLTQDKIKTHGAALOCRTTEDPNNG 360
QY      361 FRPDGTITAYRSPGAGVRLDGAOLGGEITAHFDSMLVKMTCRSDEPTAVARORAL 420
Db      361 FRPDGTITAYRSPGAGVRLDGAOLGGEITAHFDSMLVKMTCRSDEPTAVARORAL 420
QY      421 AEFTVSGVATNIGFLRALLREDEFTSKRIATGFIADHPHLLQAPPADDEGRITLDYADY 480
Db      421 AEFTVSGVATNIGFLRALLREDEFTSKRIATGFIADHPHLLQAPPADDEGRITLDYADY 480
QY      481 TVNKPGRVPRKDVAAIDIKLPNIKDLPLRGSHDRKLCQIGRAAFANDLREODALAVTDTT 540
Db      481 TVNKPGRVPRKDVAAIDIKLPNIKDLPLRGSHDRKLCQIGRAAFANDLREODALAVTDTT 540
QY      541 FRDAGSLATRYRSFALRPAEAAVAKLTPELLISVAMGATYDVMARLFEDPMRDLDE 600
Db      541 FRDAGSLATRYRSFALRPAEAAVAKLTPELLISVAMGATYDVMARLFEDPMRDLDE 600
QY      601 LREAMPNVNIOMLRGNTVGYTPYDVSCRAVFKVKAASSGVDIFRIFDALNDVSGMRPA 660
Db      601 LREAMPNVNIOMLRGNTVGYTPYDVSCRAVFKVKAASSGVDIFRIFDALNDVSGMRPA 660
QY      661 IDAVLETNAVAEVAAYSGDLSDPNEKIYTDIYIKMAEBIYKSAHILAIKDAAGLIR 720
Db      661 IDAVLETNAVAEVAAYSGDLSDPNEKIYTDIYIKMAEBIYKSAHILAIKDAAGLIR 720
QY      721 PAAVTKLVTLRREFDLPVHVTHTDYGQGLATYFAAAQAGADAVNGASAPISGTSOPS 780
Db      721 PAAVTKLVTLRREFDLPVHVTHTDYGQGLATYFAAAQAGADAVNGASAPISGTSOPS 780
QY      781 LSAIYAAFAHTRDGTGLSEAVSDLEPYWEAVRGILYLPRESGTPGFTGVYRHEIIPGQL 840
Db      781 LSAIYAAFAHTRDGTGLSEAVSDLEPYWEAVRGILYLPRESGTPGFTGVYRHEIIPGQL 840
QY      841 SNLRQAOTATLGLADREFELIEDNYAAVNEMLGRPTKVTSPSKVVGDLALHVGAGVDPADF 900
Db      841 SNLRQAOTATLGLADREFELIEDNYAAVNEMLGRPTKVTSPSKVVGDLALHVGAGVDPADF 900
QY      901 AADPOKYDIPDSVIAFLRGELGNPPGWEPELRTALLEGSRSGKAPLITEVPEEQAHDLDA 960
Db      901 AADPOKYDIPDSVIAFLRGELGNPPGWEPELRTALLEGSRSGKAPLITEVPEEQAHDLDA 960
QY      961 DDKSKERNLNLPPKPTPEEFLEHRRRGNTSALDDREFYGLVGBRETLRLDPVTRP 1020
Db      961 DDKSKERNLNLPPKPTPEEFLEHRRRGNTSALDDREFYGLVGBRETLRLDPVTRP 1020
QY      1021 LLVRLDAISEPDDKGRNVAVVNGOIRPMRVKDSVSVTATAKASSNKGHVAAPRA 1080
Db      1021 LLVRLDAISEPDDKGRNVAVVNGOIRPMRVKDSVSVTATAKASSNKGHVAAPRA 1080
QY      1081 GVVYTVVAGDEVKAGDAVAIIEAKKMEATITASVDGKIDRVVVAATRVESGDLIVVYS 1140
Db      1081 GVVYTVVAGDEVKAGDAVAIIEAKKMEATITASVDGKIDRVVVAATRVESGDLIVVYS 1140

```

RESULT 4
 AAU98053
 ID AAU98053 standard; Protein; 1140 AA.
 XX
 AC AAU98053;

```

XX      27-AUG-2002 (first entry)
DT
XX      Corynebacterium wild-type feedback-resistant pyruvate carboxylase enzyme.
DE
XX      Feedback-resistant; pyruvate carboxylase; enzyme;
XX      aspartic acid feedback inhibition resistant.
KW
XX      Corynebacterium glutamicum.
OS
XX      WO200231158-A2.
PN
XX      18-APR-2002.
PD
XX      12-OCT-2001; 2001WO-US31893.
PF
XX      13-OCT-2000; 2000US-239913P.
PR
XX      (ARCH ) ARCHER-DANIELS MIDLAND CO.
PA
XX      Hanke PD;
PI
XX      WPI; 2002-463267/49.
XX
DR      Novel mutated, feedback resistant pyruvate carboxylase enzyme
PT      polypeptide, useful for producing amino acids e.g. L-lysine,
PT      L-threonine, L-glycine, L-glutamic acid, L-proline and L-methionine and
PT      L-isoleucine
PS      Disclosure; Fig 2; 42pp; English.
XX
CC      The present invention relates to a new mutated, feedback-resistant
CC      pyruvate carboxylase enzyme. The invention is useful for producing an
CC      amino acid (e.g. L-lys, L-Thr, L-Met, L-Ile, L-Glu, L-Arg and L-Pro),
CC      by culturing a host cell in a suitable media and separating the amino
CC      acid from the medium. The vector of the invention is useful for
CC      replacement of a wild-type pyruvate carboxylase gene, with a feedback
CC      resistant pyruvate carboxylase gene, in a Corynebacterium glutamicum, by
CC      replacing a genomic copy of the wild-type pyruvate carboxylase gene with
CC      a selectable marker gene through homologous recombination to form a first
CC      recombinant strain, and replacing the selectable marker gene in the
CC      first recombinant strain, with feedback resistant pyruvate carboxylase
CC      gene through homologous recombination to form a second recombinant
CC      strain, where the homologous recombination in the above steps, occurs
CC      between the host cell and the vector. The feedback-resistant pyruvate
CC      carboxylase enzyme is resistant to feedback inhibition from aspartic
CC      acid. The present amino acid sequence represents the wild-type
CC      feedback-resistant pyruvate carboxylase enzyme of the invention.
XX
SQ      Sequence 1140 AA;

```

Query Match 100.0%; Score 5788; DB 23; Length 1140;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 MSHHTSSTLPAPFKKILVANRGEIAVARFAALETGAATVAIYPRDRGSFHRSEASEAYR 60
Db      1 MSHHTSSTLPAPKKILVANRGEIAVARFAALETGAATVAIYPRDRGSFHRSEASEAYR 60
QY      61 IGTESSPVKAYLDDDEITIGAAKKVADAIYGYGGLSENQALRCAENGITFTGPTPEV 120
Db      61 IGTESSPVKAYLDDDEITIGAAKKVADAIYGYGGLSENQALRCAENGITFTGPTPEV 120
QY      121 LDLTGDKSRAYTAAKAGLPVLAESTPSKNIDEIVKSAEGQYPIEVKAAVAGGGRMFR 180
Db      121 LDLTGDKSRAYTAAKAGLPVLAESTPSKNIDEIVKSAEGQYPIEVKAAVAGGGRMFR 180
QY      181 VASPELRLKLTASREAEAFGDAVYVERAVINPOHIEVOILGHTGEVHLYERDCS 240
Db      181 VASPELRLKLTASREAEAFGDAVYVERAVINPOHIEVOILGHTGEVHLYERDCS 240
QY      241 LORRHOKVVEIAPAOHLDPDLRDRICADAVKFCRSIGYGAGTVEFLVDEKGNHVFLEM 300
Db      241 LORRHOKVVEIAPAOHLDPDLRDRICADAVKFCRSIGYGAGTVEFLVDEKGNHVFLEM 300

```

Db 241 LQRRHQVVEIAPQHLDELRLDRICADAVKFCRSIGYGAGTVEFLVDEKGNHFIEMN 300
 Qy 301 PRIOVHTYEETVEVDLVKAQMRLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNNG 360
 Db 301 PRIOVHTYEETVEVDLVKAQMRLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNNG 360
 Qy 361 FRPDGTITAYRSPGAGVRLDGAOAGETTAHFDSMLVKMTCRGSDFETAVARAORAL 420
 Db 361 FRPDGTITAYRSPGAGVRLDGAOAGETTAHFDSMLVKMTCRGSDFETAVARAORAL 420
 Qy 421 AEFTVSGVATNIGFLRALRLREEDFTSKRIATGFIADHPHLLQAPPADDEGRILYADV 480
 Db 421 AEFTVSGVATNIGFLRALRLREEDFTSKRIATGFIADHPHLLQAPPADDEGRILYADV 480
 Qy 481 TVNKPCHGVPRKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAPARLREODALAVYDTT 540
 Db 481 TVNKPCHGVPRKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAPARLREODALAVYDTT 540
 Qy 541 FRDHOSILATRVRSFALKPAAEAVAKLPPELLSVEAMGATYDVAMRFLFEDPMRLDE 600
 Db 541 FRDHOSILATRVRSFALKPAAEAVAKLPPELLSVEAMGATYDVAMRFLFEDPMRLDE 600
 Qy 601 LREAMPNVNIQMLLRGNTGYTPYPDSVGRAFVKEAASSGVDFIRFDALNDVSQMRPA 660
 Db 601 LREAMPNVNIQMLLRGNTGYTPYPDSVGRAFVKEAASSGVDFIRFDALNDVSQMRPA 660
 Qy 661 IDAVLENTNVAEVAAMAYSGDLSDPNKKLTLDYLYLKMAEITYKSGHIIAIDMAGLLR 720
 Db 661 IDAVLENTNVAEVAAMAYSGDLSDPNKKLTLDYLYLKMAEITYKSGHIIAIDMAGLLR 720
 Qy 721 PAATKLTVTALRREFDLPVHVHTHDPAGOLATYFAAQAQADAVDASAPLSGTTQSPS 780
 Db 721 PAATKLTVTALRREFDLPVHVHTHDPAGOLATYFAAQAQADAVDASAPLSGTTQSPS 780
 Qy 781 LSAIVAAFAHTRDGTGLSEAVSDLEPYEAVRGVLYLPFESGTPGTRGVYRHEIPGGQL 840
 Db 781 LSAIVAAFAHTRDGTGLSEAVSDLEPYEAVRGVLYLPFESGTPGTRGVYRHEIPGGQL 840
 Qy 841 SNIARAOTATGLADREFLEDNYAAVNMELGRPTKVPSSKVVGDALHLVAGVDPADF 900
 Db 841 SNIARAOTATGLADREFLEDNYAAVNMELGRPTKVPSSKVVGDALHLVAGVDPADF 900
 Qy 901 AADPOKYDIPDSYIAFLRGEIAGNPGMPPEPLRTALEGRSEKAPLTEYEEEOAHIDA 960
 Db 901 AADPOKYDIPDSYIAFLRGEIAGNPGMPPEPLRTALEGRSEKAPLTEYEEEOAHIDA 960
 Qy 961 DSKERRNSLNRLLFPKPTPEEFLHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVTRP 1020
 Db 961 DSKERRNSLNRLLFPKPTPEEFLHRRRFGNTSALDDREFFYGLVEGRETLIRLPDVTRP 1020
 Qy 1021 LTVRLDAISRPDDKGMNVVAVANGOTRPMRVDRDSVESYTAATEKADSSNKGIVAAPFA 1080
 Db 1021 LTVRLDAISRPDDKGMNVVAVANGOTRPMRVDRDSVESYTAATEKADSSNKGIVAAPFA 1080
 Qy 1081 GVTVTVAEGDEKAGDAVAIIIEAMKMEATITASVDKIDRVVVPATKVEGGLIYVVS 1140
 Db 1081 GVTVTVAEGDEKAGDAVAIIIEAMKMEATITASVDKIDRVVVPATKVEGGLIYVVS 1140
 RESULT 5
 AAB01436
 ID AAB01436 standard; Protein: 1141 AA.
 AC AAB01436;
 XX 20-OCT-2000 (first entry)
 DE Pyruvate carboxylase of C. glutamicum.
 XX
 KW Pyruvate carboxylase; expression: amino acid biosynthesis; lysine;
 XX glutamic acid; oxaloacetate; fermentation; biosynthesis.
 OS Corynebacterium glutamicum.

XX
 PN W0200039305-A1.
 XX
 PD 06-JUL-2000.
 XX
 PF 23-DEC-1998; 98WO-US27301.
 XX
 PR 23-DEC-1998; 98WO-US27301.
 XX
 PA (SINS/) SINSKEY A J.
 PA (LESS/) LESSARD P A.
 PA (WILL/) WILLIS L B.
 XX
 PI Sinskey AJ, Lessard PA, Willis LB;
 XX
 DR WPI: 2000-465746/40.
 DR N-PSDB: AAA47533.
 XX
 PT Novel polynucleotides encoding Corynebacterium glutamicum pyruvate
 PT carboxylase useful for industrial fermentation processes comprises a
 PT specific nucleotide sequence
 XX
 PS Claim 3; Fig 1; 51pp; English.
 XX
 CC The pyruvate carboxylase of Corynebacterium glutamicum can be used
 CC for producing amino acids, preferably lysine and glutamic acid in
 CC industrial fermentations and for replenishing oxaloacetate consumed
 CC for biosynthesis during growth. By incorporating the pyruvate
 CC carboxylase gene in expression vectors levels of expression can be
 CC 2 - 20 fold higher than in Corynebacterium glutamicum.
 XX
 SQ Sequence 1141 AA:
 Qy 1 MSTHTSSTLPFAKKIILVANGELIVRAFALEGAATVAIYPRDGSFHRSPASAVR 60
 Db 2 VSTHTSSTLPFAKKIILVANGELIVRAFALEGAATVAIYPRDGSFHRSPASAVR 61
 Qy 61 IGTEGSPVAKYLDIDEIIGAARKKADAIYGYGFLSEMAQLARECAENGITTFGPPREV 120
 Db 62 IGTEGSPVAKYLDIDEIIGAARKKADAIYGYGFLSEMAQLARECAENGITTFGPPREV 121
 Qy 121 LDLTGDKSRVATAAKKAGLPYLAESTPSKINIDEIVKSAEGQYIFPKAAGGGGRCMRP 180
 Db 122 LDLTGDKSRVATAAKKAGLPYLAESTPSKINIDEIVKSAEGQYIFPKAAGGGGRCMRP 181
 Qy 181 VASPDDELRLKLTAEASREAEAFDGAAYVERAVINPQHIEYQILGDHTGEVYHLXERDCS 240
 Db 182 VASPDDELRLKLTAEASREAEAFDGAAYVERAVINPQHIEYQILGDHTGEVYHLXERDCS 241
 Qy 241 LQRRHQVVEIAPQHLDELRLDRICADAVKFCRSIGYGAGTVEFLVDEKGNHFIEMN 300
 Db 242 LQRRHQVVEIAPQHLDELRLDRICADAVKFCRSIGYGAGTVEFLVDEKGNHFIEMN 301
 Qy 301 PRIOVHTYEETVEVDLVKAQMRLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNNG 360
 Db 302 PRIOVHTYEETVEVDLVKAQMRLAAGATLKEGLTQDKIKTHGAALQCRITTEDPNNG 361
 Qy 361 FRPDGTITAYRSPGAGVRLDGAOAGETTAHFDSMLVKMTCRGSDFETAVARAORAL 420
 Db 362 FRPDGTITAYRSPGAGVRLDGAOAGETTAHFDSMLVKMTCRGSDFETAVARAORAL 421
 Qy 421 AEFTVSGVATNIGFLRALRLREEDFTSKRIATGFIADHPHLLQAPPADDEGRILYADV 480
 Db 422 AEFTVSGVATNIGFLRALRLREEDFTSKRIATGFIADHPHLLQAPPADDEGRILYADV 481
 Qy 481 TVNKPCHGVPRKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAPARLREODALAVYDTT 540
 Db 482 TVNKPCHGVPRKDVAAPIDKLPNIKDLPLPRGSRDLKQLGPAAPARLREODALAVYDTT 541

```
QY 541 FRDAHQSLATRVRSFALKPAEAAVAKLTPELLSVEAMGATYDVAMRFEDPWRDLDE 600
DB 542 FRDAHQSLATRVRSFALKPAEAAVAKLTPELLSVEAMGATYDVAMRFEDPWRDLDE 601
QY 601 LREAMPNVNIOMLLRGRTVGTTPYDSCRAFYKEAASSGVDIRIFDALNDVSQMPRA 660
DB 602 LREAMPNVNIOMLLRGRTVGTTPYDSCRAFYKEAASSGVDIRIFDALNDVSQMPRA 661
QY 661 IDAVLETTAAVAEVAAMAYSGDLSDPNEKLYTDYLLKMAEEIVKSGAHILAIKDMAGLLR 720
DB 662 IDAVLETTAAVAEVAAMAYSGDLSDPNEKLYTDYLLKMAEEIVKSGAHILAIKDMAGLLR 721
QY 721 PAAVTKLVLTALRREDLPVHVHTHTDAGGOLATYFAAAGADAVDASAPLSGTTSQPS 780
DB 722 PAAVTKLVLTALRREDLPVHVHTHTDAGGOLATYFAAAGADAVDASAPLSGTTSQPS 781
QY 781 LSAIYAAFAHTRRDGLSLEAVSDLEPYWEAVRGILYLPREESTPGPTGTVYRHEIPGGL 840
DB 782 LSAIYAAFAHTRRDGLSLEAVSDLEPYWEAVRGILYLPREESTPGPTGTVYRHEIPGGL 841
QY 841 SMLRQAOTATGLADRFELIEDNYAAVNEMIGRPTVTSKRYVGDALHLVGAGVDPADF 900
DB 842 SMLRQAOTATGLADRFELIEDNYAAVNEMIGRPTVTSKRYVGDALHLVGAGVDPADF 901
QY 901 AADPOKYIDIPDSVIAFLRGELGNPPGWPPEPLRTALBGRSEGKAPLTEVPEEQAHILDA 960
DB 902 AADPOKYIDIPDSVIAFLRGELGNPPGWPPEPLRTALBGRSEGKAPLTEVPEEQAHILDA 961
QY 961 DDSKRRRSINLLRPKPTPEEFLEHRRRGNTSALDDREFFGVLGEGRETLLRLDPVTRP 1020
DB 962 DDSKRRRSINLLRPKPTPEEFLEHRRRGNTSALDDREFFGVLGEGRETLLRLDPVTRP 1021
QY 1021 LLVRLDAISEPDDCKMRNVANVNGQIRPMRVDRSVESSVTATAEKADSSNKGHVAAPRA 1080
DB 1022 LLVRLDAISEPDDCKMRNVANVNGQIRPMRVDRSVESSVTATAEKADSSNKGHVAAPRA 1081
QY 1081 GVVTYTVABEGDEVKAGDAVAITTEAMKMEATTITASVDGKITDRVYVNAATKVEGGDLTVVVS 1140
DB 1082 GVVTYTVABEGDEVKAGDAVAITTEAMKMEATTITASVDGKITDRVYVNAATKVEGGDLTVVVS 1141

RESULT 6
AAM93971
ID AAM93971 standard; Protein; 1140 AA.
XX
AC AAM93971;
DT 30-JUN-1999 (first entry)
XX
DE C. glutamicum pyruvate carboxylase protein.
XX
KW Pyruvate carboxylase; amino acid production; lysine production;
KW threonine production; homoserine production; glutamate production;
KW arginine production; feed additive; condiment; pharmaceutical;
KW fine chemical; ss.
XX
OS Corynebacterium glutamicum.
XX
PN DE19831609-A1.
PD 15-APR-1999.
XX
PF 14-JUL-1998; 98DE-1031609.
XX
PR 04-OCT-1997; 97DE-1043894.
XX
PA (KERJ ) FORSCHUNGSZENTRUM JUELICH GMBH.
XX
PI Elkmanns B, Peters-Wendisch P, Sahm H;
XX
DR WP1: 1999-245521/21.
DR N-PSDB; AAX24102.
XX
```

```
PT Increasing microbial production of specific amino acids by
PT Increasing activity or expression of pyruvate carboxylase
XX
PS Claim 15; Page 11-15; 18pp; German.
XX
CC This invention describes the isolation of a pyruvate carboxylase
CC from Corynebacterium glutamicum which is used in a novel method for
CC production of lysine, threonine, homoserine, glutamate and/or arginine,
CC variously useful as feed additives, condiments, pharmaceuticals and
CC intermediates for fine chemicals. Increasing pyruvate carboxylase
CC activity increases the yield of microbial production of amino acids
CC of the aspartate and/or glutamate families, e.g. about 50% more lysine,
CC 40% more threonine and 150% more homoserine are secreted into the
CC culture medium.
XX
SQ Sequence 1140 AA;
Query Match 99.9%; Score 5782; DB 20; Length 1140;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1139; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MSTRHSSSTLLPARKKTLVANKRGIAVRAFRALFETAAIVAIYPRDRSGSFHRSFASVLR 60
DB 1 MSTRHSSSTLLPARKKTLVANKRGIAVRAFRALFETAAIVAIYPRDRSGSFHRSFASVLR 60
QY 61 IGTEGSPYKAYLDIDEIIGAANKKVKADAIYGYGGLSNAQLARCAENGITFICPTPEV 120
DB 61 IGTEGSPYKAYLDIDEIIGAANKKVKADAIYGYGGLSNAQLARCAENGITFICPTPEV 120
QY 121 LDLTGDKSRAYTAAKKAGLPVLAESTPSKNIDEIYKSAEGOTYPIFYKAVAGGGGRMF 180
DB 121 LDLTGDKSRAYTAAKKAGLPVLAESTPSKNIDEIYKSAEGOTYPIFYKAVAGGGGRMF 180
QY 181 VASPELKKLATEASREAAFGDGVYVERAVINPOIEVOIIGDTHGEVYHLYTERCS 240
DB 181 VASPELKKLATEASREAAFGDGVYVERAVINPOIEVOIIGDTHGEVYHLYTERCS 240
QY 241 LQRRHQKVEIAPAOHLDELDRICADAVKFCRSIGYGAGTVEFVLDEKGNHVFEMN 300
DB 241 LQRRHQKVEIAPAOHLDELDRICADAVKFCRSIGYGAGTVEFVLDEKGNHVFEMN 300
QY 301 PRIOVEHTVEEYVTVYDLYKAOMRLAAGATLKEIGLTQDOKIKTHGAALQCRITTEDPNG 360
DB 301 PRIOVEHTVEEYVTVYDLYKAOMRLAAGATLKEIGLTQDOKIKTHGAALQCRITTEDPNG 360
QY 361 FRPDGTITTAAYRSPGAGVRLDGAALGSETTAHFDSMLVMTCRGSPFETAVARAQRAL 420
DB 361 FRPDGTITTAAYRSPGAGVRLDGAALGSETTAHFDSMLVMTCRGSPFETAVARAQRAL 420
QY 421 AEFTVSGVATNIGFLRALIREDEFTSKRIATGFIADHPHLLQADPADDEGRILLDYADV 480
DB 421 AEFTVSGVATNIGFLRALIREDEFTSKRIATGFIADHPHLLQADPADDEGRILLDYADV 480
QY 481 TVNKPFGVPRPDVAAPIDKLPINIKDPLPRGSRRLKOLGPAARLRDRDREDAALAVDFT 540
DB 481 TVNKPFGVPRPDVAAPIDKLPINIKDPLPRGSRRLKOLGPAARLRDRDREDAALAVDFT 540
QY 541 FRDAHQSLATRVRSFALKPAEAAVAKLTPELLSVEAMGATYDVAMRFEDPWRDLDE 600
DB 541 FRDAHQSLATRVRSFALKPAEAAVAKLTPELLSVEAMGATYDVAMRFEDPWRDLDE 600
QY 601 LREAMPNVNIOMLLRGRTVGTTPYDSCRAFYKEAASSGVDIRIFDALNDVSQMPRA 660
DB 601 LREAMPNVNIOMLLRGRTVGTTPYDSCRAFYKEAASSGVDIRIFDALNDVSQMPRA 660
QY 661 IDAVLETTAAVAEVAAMAYSGDLSDPNEKLYTDYLLKMAEEIVKSGAHILAIKDMAGLLR 720
DB 661 IDAVLETTAAVAEVAAMAYSGDLSDPNEKLYTDYLLKMAEEIVKSGAHILAIKDMAGLLR 720
QY 721 PAAVTKLVLTALRREDLPVHVHTHTDAGGOLATYFAAAGADAVDASAPLSGTTSQPS 780
DB 721 PAAVTKLVLTALRREDLPVHVHTHTDAGGOLATYFAAAGADAVDASAPLSGTTSQPS 780
```

QY 781 LSAIVAAFAHTRDRTGSLAEVSDLEPYWEAVRGVLYPFESGPTGPGRYRHEIPGOL 840
D 781 LSAIVAAFAHTRDRTGSLAEVSDLEPYWEAVRGVLYPFESGPTGPGRYRHEIPGOL 840
QY 841 SNLRQAQTALGLADREFELIEDNTAAVNMELGRTKVPSSKVVGDALHLVGAGVDPAD 900
D 841 SNLRQAQTALGLADREFELIEDNTAAVNMELGRTKVPSSKVVGDALHLVGAGVDPAD 900
QY 901 AADPKYDIPDSVIAFLRGELGNPGRPEPLRFRALGEGSEGAFLTEVPEEQALDA 960
D 901 AADPKYDIPDSVIAFLRGELGNPGRPEPLRFRALGEGSEGAFLTEVPEEQALDA 960
QY 961 DSKERRNSLNRLFPKTEEFLEHRRRFGNTSALDREFFGYEGRETLIRLPDVRTP 1020
D 961 DSKERRNSLNRLFPKTEEFLEHRRRFGNTSALDREFFGYEGRETLIRLPDVRTP 1020
QY 1021 LLYRLDAISEPDDKGRNVVAVNGOIRPKRVDRSVESVTAPEKADSSNKGHVAAPE 1080
D 1021 LLYRLDAISEPDDKGRNVVAVNGOIRPKRVDRSVESVTAPEKADSSNKGHVAAPE 1080
QY 1081 GVVTVVAEGDEVKAGDAVAIIEAMKEATITASVDGKIDRVVVPATKVEGGDLIVVVS 1140
D 1081 GVVTVVAEGDEVKAGDAVAIIEAMKEATITASVDGKIDRVVVPATKVEGGDLIVVVS 1140
RESULT 7
AAG93249 standard; Protein: 1140 AA.
AC AAG93249:
XX 26-SEP-2001 (first entry)
XX C glutamicum protein fragment mutant P4585.
XX
XX Coryneform bacterium; amino acid synthesis: vitamin; saccharide;
XX organic acid synthesis; mutant; mutein.
XX Corynebacterium glutamicum.
OS Synthetic.
OS
FH Key Location/Qualifiers
FT MISC-difference 458
FT /note= "wild-type Pro substituted by Ser"
XX
PN EP1108790-A2.
XX
XX 20-JUN-2001.
XX
XX 18-DEC-2000; 2000EP-0127688.
XX
XX 16-DEC-1999; 99JP-0377484.
PR 07-APR-2000; 2000JP-0159162.
PR 03-AUG-2000; 2000JP-0280988.
XX
XX (KYOW) KYOWA HAKKO KOGYO KK.
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
XX WPI; 2001-376931/40.
XX
XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
XX
PS Claim 43; Page -: 246pp + Sequence Listing: English.
XX
XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and

CC analysing the expression profile or expression pattern of a gene derived
CC from Coryneform bacterium, and identifying a homologue of a gene derived
CC from coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a mutant protein described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
XX Sequence 1140 AA:
SQ
Query Match 99.9%; Score 5780; DB 22; Length 1140;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1139; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MSHTSSTLPFAKKILVANGELAVRAFAALETGAATVAIYPREDGSPFRSFAVA 60
D 1 MSHTSSTLPFAKKILVANGELAVRAFAALETGAATVAIYPREDGSPFRSFAVA 60
QY 61 IGTGSGVKAVIDIDELITGAKKKADAIYPGFGFSENNQLAHECAENGITFGPTPEV 120
D 61 IGTGSGVKAVIDIDELITGAKKKADAIYPGFGFSENNQLAHECAENGITFGPTPEV 120
QY 121 LDLTGDKSRVAVTAARKAGLPVLAESTPSKNIDEIVKSAEGQYPIFYKAVAGGGRGMRF 180
D 121 LDLTGDKSRVAVTAARKAGLPVLAESTPSKNIDEIVKSAEGQYPIFYKAVAGGGRGMRF 180
QY 181 VASPELRLKATEASREAEAFDGAAYVERAVINPOHIEVOILGDHTEGVHLIERDCS 240
D 181 VASPELRLKATEASREAEAFDGAAYVERAVINPOHIEVOILGDHTEGVHLIERDCS 240
QY 241 LQRRHQVVEIAPQOHIDPELRDRIKADANKFCGSIYQAGATVEFLVDEGNHVFTEM 300
D 241 LQRRHQVVEIAPQOHIDPELRDRIKADANKFCGSIYQAGATVEFLVDEGNHVFTEM 300
QY 301 PRIOVHTVEEVEVDLVKAOMLAAGATLKEGLTQDKIKTHGALQCRITTEDPNNG 360
D 301 PRIOVHTVEEVEVDLVKAOMLAAGATLKEGLTQDKIKTHGALQCRITTEDPNNG 360
QY 361 FRPDTGTTAVRSPGAGVRLDGAOLGSETTAHFDSMLVKMTCRGSDFETAVARAORAL 420
D 361 FRPDTGTTAVRSPGAGVRLDGAOLGSETTAHFDSMLVKMTCRGSDFETAVARAORAL 420
QY 421 AEFVSGVANTNIGFLRALREDEFTSKRIATGFIADHSHLQAPPADDEGRILDIYADV 480
D 421 AEFVSGVANTNIGFLRALREDEFTSKRIATGFIADHSHLQAPPADDEGRILDIYADV 480
QY 481 TVNKPGRVPRKDVAPIDKLPNIKDLPLPGSRDRLQOLGPAFARDLREODALAVYDTT 540
D 481 TVNKPGRVPRKDVAPIDKLPNIKDLPLPGSRDRLQOLGPAFARDLREODALAVYDTT 540
QY 541 FRDAHQSLLATRVRSFALKPAEAVAKLTPELLSVEAMGATYDVAMRFLFEDPWDRUDE 600
D 541 FRDAHQSLLATRVRSFALKPAEAVAKLTPELLSVEAMGATYDVAMRFLFEDPWDRUDE 600
QY 601 LREAMPVNTOMLKGNTVGYTPYDPSVCARAFKFAASSGVDFEFLDALNDVDSQMRPA 660
D 601 LREAMPVNTOMLKGNTVGYTPYDPSVCARAFKFAASSGVDFEFLDALNDVDSQMRPA 660
QY 661 IDAVLENTAVAEVAMVSGDLSDPNEKLYTLIDYYLMAEEIVSGAHITAIKMDALLR 720
D 661 IDAVLENTAVAEVAMVSGDLSDPNEKLYTLIDYYLMAEEIVSGAHITAIKMDALLR 720
QY 721 PAAVTKLVIALRREFDLPHVHTHTDAGOLATYFAAAQAGADVAGASAPLSGTSQPS 780
D 721 PAAVTKLVIALRREFDLPHVHTHTDAGOLATYFAAAQAGADVAGASAPLSGTSQPS 780
QY 781 LSAIVAAFAHTRDRTGSLAEVSDLEPYWEAVRGVLYPFESGPTGPGRYRHEIPGOL 840
D 781 LSAIVAAFAHTRDRTGSLAEVSDLEPYWEAVRGVLYPFESGPTGPGRYRHEIPGOL 840
QY 841 SNLRQAQTALGLADREFELIEDNTAAVNMELGRTKVPSSKVVGDALHLVGAGVDPAD 900

```

Db      |||
841 SMIRQAATALGLADRFELLIEDNYAAVNEMIGRFTVTPSSKVVGDALALVAGVDPADF 900
QY      AADPOKYDIPDSVIAFLRGLGNPGGWEPELRTALERSRGKAPLTPVEPEEQAHIDA 960
Db      |||
901 AADPOKYDIPDSVIAFLRGLGNPGGWEPELRTALERSRGKAPLTPVEPEEQAHIDA 960
QY      961 DDKERRNSLNLRLPEKPTPEELHRRRRRGNTSALDDREFEYGLVEGRETLRLDPVTRP 1020
Db      |||
961 DDKERRNSLNLRLPEKPTPEELHRRRRRGNTSALDDREFEYGLVEGRETLRLDPVTRP 1020
QY      1021 LLVRLDAISEPDDKGMNVNANVQIRPMRVDRDSVESVTATAEKADSSNKGHYAAPPA 1080
Db      |||
1021 LLVRLDAISEPDDKGMNVNANVQIRPMRVDRDSVESVTATAEKADSSNKGHYAAPPA 1080
QY      1081 GVVYTVVAGDEVKAGDAVAITEAKMKMETITASVDGKIDRVVPAATVVEGGDLIVVVS 1140
Db      |||
1081 GVVYTVVAGDEVKAGDAVAITEAKMKMETITASVDGKIDRVVPAATVVEGGDLIVVVS 1140

RESULT 8
AAU98052
ID      AAU98052 standard; Protein; 1140 AA.
AC      AAU98052;
DT      27-AUG-2002 (first entry)
XX      Corynebacterium mutant feedback-resistant pyruvate carboxylase enzyme.
DE      Corynebacterium mutant feedback-resistant pyruvate carboxylase enzyme.
XX      Feedback-resistant; pyruvate carboxylase; enzyme;
KW      aspartic acid feedback inhibition resistant; mutant; mutein.
OS      Corynebacterium glutamicum.
XX      Synthetic.
FH      Key
FH      Location/Qualifiers
FT      Misc-difference 1 /note= "Wild-type Met substituted by Val"
FT      Misc-difference 153 /note= "Wild-type Glu substituted by Asp"
FT      Misc-difference 182 /note= "Wild-type Ala substituted by Ser"
FT      Misc-difference 206 /note= "Wild-type Ala substituted by Ser"
FT      Misc-difference 227 /note= "Wild-type Ala substituted by Ser"
FT      Misc-difference 455 /note= "Wild-type His substituted by Arg"
FT      Region /note= "Wild-type Ala substituted by Gly"
FT      Misc-difference 1116 /note= "Specifically claimed in claim 18"
FT      Misc-difference /note= "Wild-type Asp substituted by Glu"
PN      WO200231158-A2.
XX      18-APR-2002.
PD      12-OCT-2001; 2001WO-US31893.
PF      13-OCT-2000; 2000US-239913P.
PR      (ARCH ) ARCHER-DANIELS MIDLAND CO.
PA      Hanke PD;
XX      WPI: 2002-463267/49.
XX      Novel mutated, feedback resistant pyruvate carboxylase enzyme
PT      polypeptide, useful for producing amino acids e.g. L-lysine,
PT      L-threonine, L-glycine, L-glutamic acid, L-proline and L-methionine and
PT      L-isoleucine
XX

```

```

PS      Claim 1: Page -: 42pp; English.
XX      The present invention relates to a new mutated, feedback-resistant
CC      pyruvate carboxylase enzyme. The invention is useful for producing an
CC      amino acid (e.g. L-Lys, L-Thr, L-Met, L-Ile, L-Glu, L-Arg and L-Pro),
CC      by culturing a host cell in a suitable media and separating the amino
CC      acid from the medium. The vector of the invention is useful for
CC      replacement of a wild-type pyruvate carboxylase gene, with a feedback
CC      resistant pyruvate carboxylase gene, in a Corynebacterium glutamicum, by
CC      replacing a genomic copy of the wild-type pyruvate carboxylase gene with
CC      a selectable marker gene through homologous recombination to form a first
CC      recombinant strain, and replacing the selectable marker gene in the
CC      first recombinant strain, with feedback resistant pyruvate carboxylase
CC      gene through homologous recombination to form a second recombinant
CC      strain, where the homologous recombination in the above steps, occurs
CC      between the host cell and the vector. The feedback-resistant pyruvate
CC      carboxylase enzyme is resistant to feedback inhibition from aspartic
CC      acid. The present amino acid sequence represents the mutant
CC      feedback-resistant pyruvate carboxylase enzyme of the invention.
CC      Note: The present sequence is not shown in the specification but is
CC      derived from the wild-type feedback-resistant pyruvate carboxylase
CC      enzyme (AAU98053) given in figure 2 of the specification.
XX      Sequence 1140 AA;

```

```

Query Match          99.5%; Score 5759; DB 23; Length 1140;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1133; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      1 MSTRSSILPFAFKKLLVNRGEIYAVRAFALETCATVAIYIPREDRGSFHRFSASEVR 60
Db      1 VSTRSSILPFAFKKLLVNRGEIYAVRAFALETCATVAIYIPREDRGSFHRFSASEVR 60
QY      61 IGTEGSPYKAYLDDEIIGAAKKVYKADLYGCGYGLSNAOLARCAANGTIFGPEYEV 120
Db      61 IGTEGSPYKAYLDDEIIGAAKKVYKADLYGCGYGLSNAOLARCAANGTIFGPEYEV 120
QY      121 LDLTGDKSRAYTAARKAGLPVLAESTPSKNIDETVYKSAEGQTYPIFYKAVAGGGGRGMR 180
Db      121 LDLTGDKSRAYTAARKAGLPVLAESTPSKNIDETVYKSAEGQTYPIFYKAVAGGGGRGMR 180
QY      181 VASPELKLKLTASREAEAFGCGAVYVERAVINPQITEVQIIGDRTGEVHHLYERCCS 240
Db      181 VASPELKLKLTASREAEAFGCGAVYVERAVINPQITEVQIIGDRTGEVHHLYERCCS 240
QY      241 LORRQKVEIAPAOHIDPELRICADAVFCRSIGYOGATVEFLVDEKGNHVFIDMN 300
Db      241 LORRQKVEIAPAOHIDPELRICADAVFCRSIGYOGATVEFLVDEKGNHVFIDMN 300
QY      301 PRIOVEHTVTEBEVTVLKAQMRLAAGATLEKELGTODKIKTHGAALQCRITTEDPNNG 360
Db      301 PRIOVEHTVTEBEVTVLKAQMRLAAGATLEKELGTODKIKTHGAALQCRITTEDPNNG 360
QY      361 FRPDGTITAYRSRSGAGVRLDGAALQGETTANFDSMLVMTKRGSPFETAVARAQAL 420
Db      361 FRPDGTITAYRSRSGAGVRLDGAALQGETTANFDSMLVMTKRGSPFETAVARAQAL 420
QY      421 AEFYSSGATNIGFLRALREDEFTSKRIATGFTADHGHLLQAPPADDEQGRITIDYADV 480
Db      421 AEFYSSGATNIGFLRALREDEFTSKRIATGFTADHGHLLQAPPADDEQGRITIDYADV 480
QY      481 TVNKPVGVRPDVAPIDKLPINIKDPLPRGSSRDRLKQLGPAARLDREODALAVYDTT 540
Db      481 TVNKPVGVRPDVAPIDKLPINIKDPLPRGSSRDRLKQLGPAARLDREODALAVYDTT 540
QY      541 FRDAHQSILATRVVSFALKPAEAVAKITPELLSYEAMGATTVYANRFLFEDPDWRIDE 600
Db      541 FRDAHQSILATRVVSFALKPAEAVAKITPELLSYEAMGATTVYANRFLFEDPDWRIDE 600
QY      601 LREAMPNVNIDMLRGRTVGYTPYPSVCAPFAKFEAASSGVDFRIFDALNDVDSORPA 660
Db      601 LREAMPNVNIDMLRGRTVGYTPYPSVCAPFAKFEAASSGVDFRIFDALNDVDSORPA 660

```


QY 661 IDAVLENTAVAEVAMAYSGDLSDPNEKLTLDLYLKMAEEIYKGAHILAIKMAQILLR 720
 DB 661 IDAVLENTAVAEVAMAYSGDLSDPNEKLTLDLYLKMAEEIYKGAHILAIKMAQILLR 720
 QY 721 PAAVTKLVTLRREFDLPVHVTHTDAGGLATYFAAAQAGADAVDASAPLSTGTSOPS 780
 DB 721 PAAVTKLVTLRREFDLPVHVTHTDAGGLATYFAAAQAGADAVDASAPLSTGTSOPS 780
 QY 781 LSAIVAAFAHTRRDGTGLSLAVSDLEPYWEAVRGLYLPESGTPGPRGYRHHIPGOL 840
 DB 781 LSAIVAAFAHTRRDGTGLSLAVSDLEPYWEAVRGLYLPESGTPGPRGYRHHIPGOL 840
 QY 841 SNLRQAQTALGLADREFLIEDNTAAVNMELGRPTKVPSSKVVGDALHLVAGADPADF 900
 DB 841 SNLRQAQTALGLADREFLIEDNTAAVNMELGRPTKVPSSKVVGDALHLVAGADPADF 900
 QY 901 AADPOKXIDIPSVIAFLRGLGNPPGMPPLRTALRSEKAPLLEVEPEEQAHILDA 960
 DB 901 AADPOKXIDIPSVIAFLRGLGNPPGMPPLRTALRSEKAPLLEVEPEEQAHILDA 960
 QY 961 DDKERNSLNRLLFPKPTSEFLEHRRRGNTSLADREFFGYLVGREGRETLIRLPDVRTP 1020
 DB 961 DDKERNSLNRLLFPKPTSEFLEHRRRGNTSLADREFFGYLVGREGRETLIRLPDVRTP 1020
 QY 1021 LTVRLDAISEPDDKGMNVAVANVGOTRPMKRVDRSVESTYATPEKADSSNKGHVAAPEFA 1080
 DB 1021 LTVRLDAISEPDDKGMNVAVANVGOTRPMKRVDRSVESTYATPEKADSSNKGHVAAPEFA 1080
 QY 1081 GVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDKIDRVVVPAAATKEGGDLIVVS 1140
 DB 1081 GVTVTVAEGDEVKAGDAVAIIEMKMEATITASVDKIDRVVVPAAATKEGGDLIVVS 1140

RESULT 9
 AAU98050 standard; Protein: 1157 AA.

AC AAU98050:
 DT 27-AUG-2002 (first entry)
 XX Corynebacterium feedback-resistant pyruvate carboxylase enzyme.
 DE Corynebacterium feedback-resistant pyruvate carboxylase enzyme.
 KM Feedback-resistant: pyruvate carboxylase; enzyme:
 KM aspartic acid feedback inhibition resistant.
 XX Corynebacterium glutamicum.
 OS
 FH Key Location/Qualifiers
 FT Region 1..18
 FT Region /note- "Specifically claimed in claim 18"
 FT Region 164..176
 FT Region /note- "Specifically claimed in claim 18"
 FT Region 193..205
 FT Region /note- "Specifically claimed in claim 18"
 FT Region 217..229
 FT Region /note- "Specifically claimed in claim 18"
 FT Region 238..250
 FT Region /note- "Specifically claimed in claim 18"
 FT Region 466..478
 FT Region /note- "Specifically claimed in claim 18"
 XX
 PN WO200231158-A2.
 XX
 PD 18-APR-2002.
 XX
 PF 12-OCT-2001; 2001WO-US31893.
 XX
 PR 13-OCT-2000; 2000US-239913P.
 XX
 PA (ARCH) ARCHER-DANIELS MIDLAND CO.
 XX
 PI Hanke PD;

XX
 DR WPI: 2002-463267/49.
 DR N-PSDB: ABK52832.
 XX
 PT Novel mutated, feedback resistant pyruvate carboxylase enzyme
 PT polypeptide, useful for producing amino acids e.g. L-lysine,
 PT L-threonine, L-glycine, L-glutamic acid, L-proline and L-methionine and
 PT L-isoleucine
 XX
 PS Claim 17, Fig 1: 42pp; English.
 XX
 PS The present invention relates to a new mutated, feedback-resistant
 CC pyruvate carboxylase enzyme. The invention is useful for producing an
 CC amino acid (e.g. L-lys, L-thr, L-Met, L-Ile, L-Glu, L-Arg and L-Pro),
 CC by culturing a host cell in a suitable media and separating the amino
 CC acid from the medium. The vector of the invention is useful for
 CC replacement of a wild-type pyruvate carboxylase gene, with a feedback
 CC resistant pyruvate carboxylase gene, in a Corynebacterium glutamicum,
 CC replacing a genomic copy of the wild-type pyruvate carboxylase gene with
 CC a selectable marker gene through homologous recombination to form a first
 CC recombinant strain, and replacing the selectable marker gene in the
 CC first recombinant strain, with feedback resistant pyruvate carboxylase
 CC gene through homologous recombination to form a second recombinant
 CC strain, where the homologous recombination in the above steps, occurs
 CC between the host cell and the vector. The feedback-resistant pyruvate
 CC carboxylase enzyme is resistant to feedback inhibition from aspartic
 CC acid. The present amino acid sequence represents the feedback-resistant
 CC pyruvate carboxylase enzyme of the invention.
 CC
 XX
 SQ Sequence 1157 AA:
 QY 1 MSTHTSSLPFAKKILVANRGEIIVRAFLAETGAATVAIYPREDGSGFRSPASEAVR 60
 DB 18 VSTHTSSLPFAKKILVANRGEIIVRAFLAETGAATVAIYPREDGSGFRSPASEAVR 77
 QY 61 IGTGSPYKAYLDIDEITIGAAKKKADAIYGYGFLSENOALRECAENGTFGTPPEV 120
 DB 78 IGTGSPYKAYLDIDEITIGAAKKKADAIYGYGFLSENOALRECAENGTFGTPPEV 137
 QY 121 LDLTGDSRAVTAARKKGLPLAESTPSKNDLTVKSEGGTYPIFKAVAGGGGGRGRF 180
 DB 138 LDLTGDSRAVTAARKKGLPLAESTPSKNDLTVKSEGGTYPIFKAVAGGGGGRGRF 197
 QY 181 VASPELRLKLTAEASREAEAFDGAAYVERAVINPOHIEVOILGDHTEGVNHYERDCS 240
 DB 198 VASPELRLKLTAEASREAEAFDGAAYVERAVINPOHIEVOILGDHTEGVNHYERDCS 257
 QY 241 LQRHOKVVEIAPAOHLDELRLRICADAVKFCISIGOGAGTYEFLVDEKGNHVFTEBN 300
 DB 258 LQRHOKVVEIAPAOHLDELRLRICADAVKFCISIGOGAGTYEFLVDEKGNHVFTEBN 317
 QY 301 PRIOVEHTVEEVEVDLVKQMLAGATLKEJLGTODKTKTGALQOCITTEDPNNG 360
 DB 318 PRIOVEHTVEEVEVDLVKQMLAGATLKEJLGTODKTKTGALQOCITTEDPNNG 377
 QY 361 FRPDGTITAVRSFGAGVRLDGAOAGETTAHFDSMLVMTGRSGSFETAVARAORAL 420
 DB 378 FRPDGTITAVRSFGAGVRLDGAOAGETTAHFDSMLVMTGRSGSFETAVARAORAL 437
 QY 421 AEFVSGVAINIGFLRALREEDFTSKRIATGFTADHPHLQAPPADDEGRITLDYADV 480
 DB 438 AEFVSGVAINIGFLRALREEDFTSKRIATGFTADHPHLQAPPADDEGRITLDYADV 497
 QY 481 TVNKPBGVAPKDVAPIDKLPINIKDLP.PRGSSRDRLQOLGPAARFARDREDDALAYDVT 540
 DB 498 TVNKPBGVAPKDVAPIDKLPINIKDLP.PRGSSRDRLQOLGPAARFARDREDDALAYDVT 557
 QY 541 FRDAHOSLTAIRVSFALKPAAEAVNAKLTPELLSVEAMGATYDVANRFLFEDPMRDLDE 600

Db 558 FRDAHOSLATTRRSALRKAFAEVAKLPELLSVEAWGATYVAMRFLFEDPMDRLDE 617
QY 601 LREAMPNVNIOMLLRGNTVGYTPYDSCRAFEVKEAASSGVDFIRIFDALNDVSOBRPA 660
Db 618 LREAMPNVNIOMLLRGNTVGYTPYDSCRAFEVKEAASSGVDFIRIFDALNDVSOBRPA 677
QY 661 IDAVLETNTNVAEVAAMAYSGDLSDPNEKLYTLDYILKMAEYIKSAHILAIKDMAGLLR 720
Db 678 IDAVLETNTNVAEVAAMAYSGDLSDPNEKLYTLDYILKMAEYIKSAHILAIKDMAGLLR 737
QY 721 PAAVTLVTALRREFLPVHVHTHDFAAGOLATYFAAAGADAVGASAPLSGTSOPS 780
Db 738 PAAVTLVTALRREFLPVHVHTHDFAAGOLATYFAAAGADAVGASAPLSGTSOPS 797
QY 781 LSAIVAAFAHTRRDGTLSLEAVSDLEPYWEAVRGLYLPESGTPGPTGRVREIIPGQOL 840
Db 798 LSAIVAAFAHTRRDGTLSLEAVSDLEPYWEAVRGLYLPESGTPGPTGRVREIIPGQOL 857
QY 841 SNIRAAQATALGLADRELIEDNYAAVNEMLGRPTKVTSPSSKVVGDIALHLVGAGVDPADP 900
Db 858 SNIRAAQATALGLADRELIEDNYAAVNEMLGRPTKVTSPSSKVVGDIALHLVGAGVDPADP 917
QY 901 AADPOKYDIPDSVIATFRLGELGNPGWPEPLRTALRGRSSEKAPLITEVPEEQAHLDA 960
Db 918 AADPOKYDIPDSVIATFRLGELGNPGWPEPLRTALRGRSSEKAPLITEVPEEQAHLDA 977
QY 961 DSKERRNSILNRLLEPKPTPEEFLEHRRRREGNTSALDDREFEYGLVGEARETLIRLPDYRTP 1020
Db 978 DSKERRNSILNRLLEPKPTPEEFLEHRRRREGNTSALDDREFEYGLVGEARETLIRLPDYRTP 1037
QY 1021 LTVRLDAISEPDDKGRNVYANVNGOIRPMRYRDSVESVYATAEKADSSNKGHVAAPPA 1080
Db 1038 LTVRLDAISEPDDKGRNVYANVNGOIRPMRYRDSVESVYATAEKADSSNKGHVAAPPA 1097
QY 1081 GVVTVTVABGDEYKAGDAVAIIEAMKMEATITASVYDGIIDRVVVPATVEGGDLIVVVS 1140
Db 1098 GVVTVTVABGDEYKAGDAVAIIEAMKMEATITASVYDGIIDRVVVPATVEGGDLIVVVS 1157

RESULT 10
AAB83180
ID AAB83180 standard; Protein; 1139 AA.
XX
AC AAB83180;
09-JUL-2001 (first entry)
DE Corynebacterium thermoaminogenes pc protein.
XX
KW Corynebacterium; thermophilic; amino acid biosynthesis; enzyme;
KW thermotolerant; acea; accBC; dtsR1; dtsR2; pfk; scrB; gluaBCD;
XX pdha; pc; ppc; acn; 1cd; 1pd; odha.
OS Corynebacterium thermoaminogenes.
XX
PN WO200125447-A1.
12-APR-2001.
PD
XX
PF 04-OCT-2000; 2000WO-JP06913.
XX
PR 04-OCT-1999; 99TP-0282716.
PR 01-NOV-1999; 99OP-0311147.
PR 21-APR-2000; 2000JP-0120687.
XX
PA (AJIN) AJINOMOTO CO INC.
XX
PI Hirano S, Nonaka G, Matsuzaki Y, Akiyoshi N, Nakamura K, kimura E;
PI Osumi T, Matsui K, Kawahara Y, Kurahashi O, Nakamatsu T;
PI Sugimoto S;
XX
XX WPI: 2001-300170/31.
DR N-PSDB; AAF87437.

XX
PT Proteins and their DNA useful for microbial production of L-amino acids
PT
XX
XX
PS Claim 9; Page 132-135; 215pp; Japanese.
XX
CC The present sequence is provided in a specification relating to genes
CC encoding thermophilic amino acid biosynthesis system enzymes of
CC the thermotolerant bacterium Corynebacterium thermoaminogenes.
CC The novel proteins retain at least 30% isocitrate lyase activity
CC after heating at 500C for 5 minutes. DNA fragments encoding the
CC enzymes were isolated from a Corynebacterium thermoaminogenes
CC chromosomal DNA plasmid library by PCR. The DNA may be used for
CC developing strains of amino acid producing microorganisms.
XX
SQ Sequence 1139 AA;
Query Match 91.6%; Score 5302.5; DB 22; Length 1139;
Best local Similarity 91.0%; Pred. No. 0;
Matches 1036; Conservative 43; Mismatches 58; Indels 1; Gaps 1;
QY 3 TTTSSTLPFAFKKILVANRGEIVNRAFRALLETGAATVAIYPREDGSHRSFASAVRIG 62
Db 3 TTTPSTLPFAFKKILVANRGEIVNRAFRALLETGAATVAIYPREDGSHRSFASAVRIG 62
QY 63 TEGSPVKATLIDIDEITIGAARKKADAIYPGYGFLSENNOLARECAENGITFETPTPEYVD 122
Db 63 TEGSPVKATLIDIDEITIGAARKKADAIYPGYGFLSENNOLARECAENGITFETPTPEYVD 122
QY 123 LTGDKSRAYTAARKKAGLPVLAESTPSKINIDEIVKSAEGQYPIFKAVAVAGGGGRMRFA 182
Db 123 LTGDKSRAYTAARKKAGLPVLAESTPSKINIDEIVKSAEGQYPIFKAVAVAGGGGRMRFA 182
QY 123 LTGDKSKKAVSAKKAGLPVLAESTPSKINIDEIVKSAEGQYPIFKAVAVAGGGGRMRFA 182
Db 123 LTGDKSKKAVSAKKAGLPVLAESTPSKINIDEIVKSAEGQYPIFKAVAVAGGGGRMRFA 182
QY 183 SPDELRLKATESSREARAEAFGAGAYVERAVNPNHIEVQLIGHTGEVNHLYEEDCSLQ 242
Db 183 SPDELRLKATESSREARAEAFGAGAYVERAVNPNHIEVQLIGHTGEVNHLYEEDCSLQ 242
QY 183 KREDLREAREASREAEAFGAGAYVERAVNPNHIEVQLIGHTGEVNHLYEEDCSLQ 242
Db 183 KREDLREAREASREAEAFGAGAYVERAVNPNHIEVQLIGHTGEVNHLYEEDCSLQ 242
QY 243 RRHQKVEIAPAOHIDPELRDRICADAVKFCISIGOGAGVVEFLVDEKGNHVFEMPR 302
Db 243 RRHQKVEIAPAOHIDPELRDRICADAVKFCISIGOGAGVVEFLVDEKGNHVFEMPR 302
QY 303 IOVEHTVTEVTVVDLVKAOMRLAAGATLKEIGLQDOKITTHGAALQCRITTEDPNSNFR 362
Db 303 IOVEHTVTEVTVVDLVKAOMRLAAGATLKEIGLQDOKITTHGAALQCRITTEDPNSNFR 362
QY 363 PDGTGTTAYRSPGCGAVRLDGAOLGSETAHPSMVLKMKCRGSDFTAVASRAQRALAE 422
Db 363 PDGTGTTAYRSPGCGAVRLDGAOLGSETAHPSMVLKMKCRGSDFTAVASRAQRALAE 422
QY 423 FTVSGVATNIGFLRALRREDFTSKRIATGFIADPHLLQAPPADDEGRIIDYLAADV 482
Db 423 FTVSGVATNIGFLRALRREDFTSKRIATGFIADPHLLQAPPADDEGRIIDYLAADV 482
QY 483 NKPHGVRKDYAAPIDKLPNTKIDLPPLPGSGDRKLKOLGPAAPANDLEQDALAVDTYFR 542
Db 483 NKPHGVRKDYAAPIDKLPNTKIDLPPLPGSGDRKLKOLGPAAPANDLEQDALAVDTYFR 542
QY 543 DAHOSLATTRRSFALRKAFAEVAKLPELLSVEAWGATYVAMRFLFEDPMDRLDELR 602
Db 543 DAHOSLATTRRSFALRKAFAEVAKLPELLSVEAWGATYVAMRFLFEDPMDRLDELR 602
QY 603 EAMPNVNIOMLLRGNTVGYTPYDSCRAFEVKEAASSGVDFIRIFDALNDVSOBRPAID 662
Db 603 EAMPNVNIOMLLRGNTVGYTPYDSCRAFEVKEAASSGVDFIRIFDALNDVSOBRPAID 662
QY 662 AVLETGTSVAAEVAAMAYSGDLSNPGEKLYTLDYILNLAEQIYDSGAHILAIKDMAGLLRRA 721
Db 662 AVLETGTSVAAEVAAMAYSGDLSNPGEKLYTLDYILNLAEQIYDSGAHILAIKDMAGLLRRA 721
QY 723 AVTKLVTALRREFLPVHVHTHDFAAGOLATYFAAAGADAVGASAPLSGTSOPS 782
Db 723 AVTKLVTALRREFLPVHVHTHDFAAGOLATYFAAAGADAVGASAPLSGTSOPS 782
QY 722 AAPKLVTLALRREFLPVHVHTHDFAAGOLATYFAAAGADAVGASAPLSGTSOPS 781
Db 722 AAPKLVTLALRREFLPVHVHTHDFAAGOLATYFAAAGADAVGASAPLSGTSOPS 781

```
Oy 783 AIVAFATRRDRLTSLAEVSDLEPYMEAVRGVLYLPFESGTPGPTGRVYRHEITPGCOLSN 842
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 782 ALVAFATRRDRLTSLAEVSDLEPYMEAVRGVLYLPFESGTPGPTGRVYRHEITPGCOLSN 841
Oy 843 LRAQATLGLADREFELIEDNYAAVNMELGRPTKYTPSSKVVGDLLALHLVGAVDPADFAA 902
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 842 LRAQAVLGLADREFELIEDYAAVNMELGRPTKYTPSSKVVGDLLALHLVGAVSPEDFAA 901
Oy 903 DPQKYDIPDSYIALRGELNPGPGWPEPLTRALEGSEKAPLTYEPEEOHLLADD 962
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 902 DPQKYDIPDSYIALRGELNPGPGWPEPLTRALEGSEKAPLTYEPEEOHLLADD 961
Oy 963 SKERRNSLNLRLFPKPEEFLHRRRRCNTSALDREFEYGLVGRETLILPUPRPLL 1022
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 962 SAERGLNLRLFPKPEEFLHRRRRCNTSALDREFEYGLVGRETLILPUPRPLV 1021
Oy 1023 VRLDAISEPDDKGRNNVAVNNGQIRPMRVDRSVEVYATAEKADSSNKGHVAPAGV 1082
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1022 VRLDAVSEPDKGRNNVAVNNGQIRPIKVRDSVEVYATAEKADATNKGHVAPAGV 1081
Oy 1083 VTYVVAEGDEYKADAVAIIEAMKMEATITASVCGKIDRVVVPATKVEGGDLIVVVS 1140
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1082 VTYVVAEGDEIKADDAVAIIIEAMKMEATITAPVDGVIDRVVVPATKVEGGDLIVVVS 1139

RESULT 11
AAB79302
ID AAB79302 standard; Protein: 532 AA.
AC
XX AAB79302;
XX
XX 30-APR-2001 (first entry)
XX
XX Corynebacterium glutamicum SMP protein sequence SEQ ID NO:120.
XX
XX Corynebacterium glutamicum; carbon metabolism and energy production;
XX SMP protein; sugar metabolism and oxidative phosphorylation protein;
XX fine chemical production; organic acid; pyrimidine base; nucleoside;
XX nonproteinogenic amino acid; purine base; proteinogenic amino acid;
XX nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
XX carboxylate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
XX diagnosis; Corynebacterium diptheriae; evolutionary study.
XX
XX Corynebacterium glutamicum.
XX
XX WO200100844-A2.
XX
XX 04-JAN-2001.
XX
XX 23-JUN-2000; 2000WO-IB00943.
XX
XX 25-JUN-1999; 99US-0141031.
XX 08-JUL-1999; 99DE-1031412.
XX 08-JUL-1999; 99DE-1031413.
XX 08-JUL-1999; 99DE-1031419.
XX 08-JUL-1999; 99DE-1031420.
XX 08-JUL-1999; 99DE-1031424.
XX 08-JUL-1999; 99DE-1031428.
XX 08-JUL-1999; 99DE-1031431.
XX 08-JUL-1999; 99DE-1031433.
XX 08-JUL-1999; 99DE-1031434.
XX 08-JUL-1999; 99DE-1031510.
XX 08-JUL-1999; 99DE-1031516.
XX 08-JUL-1999; 99DE-1031562.
XX 08-JUL-1999; 99DE-1031634.
XX 09-JUL-1999; 99DE-1032180.
XX 09-JUL-1999; 99DE-1032227.
XX 09-JUL-1999; 99DE-1032230.
XX 09-JUL-1999; 99US-0143208.
XX 14-JUL-1999; 99DE-1032924.
XX 14-JUL-1999; 99DE-1032973.
XX 14-JUL-1999; 99DE-1033005.
XX 27-AUG-1999; 99DE-1040765.
XX 31-AUG-1999; 99US-0151572.
```

```
PR 03-SEP-1999; 99DE-1042076.
PR 03-SEP-1999; 99DE-1042079.
PR 03-SEP-1999; 99DE-1042086.
PR 03-SEP-1999; 99DE-1042087.
PR 03-SEP-1999; 99DE-1042088.
PR 03-SEP-1999; 99DE-1042095.
PR 03-SEP-1999; 99DE-1042123.
PR 03-SEP-1999; 99DE-1042125.
PA (BADI ) BASF AG.
PI Pompejus M, Kroegeer B, Schroeder H, Zelder O, Haberhauer G;
DR WPI: 2001-061975/07.
XX N-PSDB: AAF71419.
XX
XX New isolated Corynebacterium glutamicum nucleic acid encoding a sugar
PT metabolism and oxidative phosphorylation protein for production or
PT modulation of production of fine chemicals e.g. amino acids,
PT carbohydrates or enzymes -
XX
XX Claim 20: Page 319-320; 1246pp: English.
PS
XX AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar
CC metabolism and oxidative phosphorylation (SMP) proteins given in
CC AAB79243 to AAB 79633 which are involved in carbon metabolism and
CC energy production. The C. glutamicum SMP gene can be used in vectors
CC (II) for expression in host cells and production or modulation of
CC production of fine chemicals, such as, an organic acid, a proteinogenic
CC or nonproteinogenic amino acid (preferred), a purine or pyrimidine base,
CC a nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty
CC acid, a diol, a carbohydrate, an aromatic compound, a vitamin, a
CC cofactor, a polyketide, or an enzyme. The presence of (I) or SMP proteins
CC (II) encoded by them are used for diagnosing the presence or activity of
CC Corynebacterium diptheriae in a subject. (I), (II), (III) or host cells
CC containing them are used to map genomes of organisms related to
CC C. glutamicum, identify and localise C. glutamicum sequences of interest,
CC in evolutionary studies, in determining SMP protein regions required
CC for function, in modulating SMP protein activity, in modulating the
CC metabolism of sugars, and in modulating high-energy molecule production
CC in a cell (i.e. ATP, NADPH).
XX
XX Sequence 532 AA;
SQ
Query Match 45.3%; Score 2621; DB 22; Length 532;
Best Local Similarity 99.8%; Pred. No. 2.6e-184;
Matches 514; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Oy 1 MSTHTSSSTLPAPFKILVANGELIVRAFRALLETGAATVATYPREDRGSFRRFASFAVR 60
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 18 VSTRHTSSSTLPAPFKILVANKRELIVRAFRALLETGAATVATYPREDRGSFRRFASFAVR 77
Oy 61 IGTGSPYKAYLDIDEIIGAARKKADATYPGYGLSENAQLARECAENGITFTGPPREV 120
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 78 IGTGSPYKAYLDIDEIIGAARKKADATYPGYGLSENAQLARECAENGITFTGPPREV 137
Oy 121 LDLTGDKSRANTAAKAGLPVLAESTPEKNIDEIYKSAEGOTYPLFYAAVAGGGGRGRF 180
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 138 LDLTGDKSRANTAAKAGLPVLAESTPEKNIDEIYKSAEGOTYPLFYAAVAGGGGRGRF 197
Oy 181 VASPELKLKLTAEASREAEAFGCGAVYVERAVINPOIIEVQIIGDHGEVYHLYERCS 240
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 198 VASPELKLKLTAEASREAEAFGCGAVYVERAVINPOIIEVQIIGDHGEVYHLYERCS 257
Oy 241 LQRRHOKVETAPAOHLDPELRDRICADAVKFCRSIGYGAGTYEFLVDKGNHVFTEMN 300
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 258 LQRRHOKVETAPAOHLDPELRDRICADAVKFCRSIGYGAGTYEFLVDKGNHVFTEMN 317
Oy 301 PRIOVERTVTEVEVDLVKAKQMLAAGATLKEGLTODKTKTGAALQCTTTEDPNC 360
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 318 PRIOVERTVTEVEVDLVKAKQMLAAGATLKEGLTODKTKTGAALQCTTTEDPNC 377
Oy 361 FRPDGTITAVRSPGAGVRLDGAAGLGETTAHFDSMLVKMTCRGSDFEFAVARAQRAL 420
```

|||||
Db 378 FRPDTGTTIAYRSPGGAGVRLDGAAGGSEITAHFMSMLVKMTCRGSDFTAVARAQRAL 437
QY 421 AEFVSVGATNIGFLRALLREEDFTSKRIATGFIADPHILOAPPADEGRILIDYADV 480
Db 438 AEFVSVGATNIGFLRALLREEDFTSKRIATGFIADPHILOAPPADEGRILIDYADV 497
QY 481 TVNKPFGVPRKDVAAPIDKLPIKIDLPPLRGSRDR 515
Db 498 TVNKPFGVPRKDVAAPIDKLPIKIDLPPLRGSRDR 532
RESULT 12
AAU00511
ID AAU00511 standard; Protein; 1148 AA.
XX
AC AAU00511;
XX
XX 07-SEP-2001 (first entry)
XX
DE Bacillus subtilis pyruvate carboxylase enzyme A.
XX
XX Pyruvate carboxylase A; pycA; oxaloacetate; Escherichia;
KM L-amino acid production; fermentation.
XX
OS Bacillus subtilis strain 168.
XX
PN EPI092776-A1.
XX
PD 18-APR-2001.
XX
PE 05-OCT-2000; 2000EP-0121763.
XX
PR 14-OCT-1999; 99RU-0121636.
XX
PA (AJIN) AJINOMOTO CO INC.
PI Gasyatlier MM, Kozlov VI, Plitsyn LR, Altman IB, Voroshilova EB;
PI Iomantas YAV, Yampolskaya TA;
XX
XX WPI; 2001-309819/33.
DR N-PSDB; AAS01509.
XX
XX New bacterium from the genus Escherichia containing a gene encoding for
PT pyruvate carboxylase useful for producing higher concentrations of
PT L-amino acids
XX
XX Claim 4; Page 17-21; 28pp; English.
XX
XX The present sequence represents Bacillus subtilis pyruvate
CC carboxylase enzyme A. The pycA gene encodes for pyruvate carboxylase
CC which catalyses the carboxylation of pyruvate to form oxaloacetate.
CC Transformation of the Bacillus subtilis pycA gene into a bacterium
CC belonging to the genus Escherichia results in the bacterium showing
CC L-amino acid productivity. The invention provides a novel method for
CC producing an L-amino acid by fermentation. The method involves
CC culturing the bacterium in a medium and producing and accumulating
CC the L-amino acid in the medium. The new bacterium harbouring the gene
CC coding for an enzyme having pyruvate carboxylase activity is useful
CC for producing higher concentrations of L-amino acids in vitro than
CC prior art.
XX
SQ Sequence 1148 AA;
Query Match 43.6%; Score 2524; DB 22; Length 1148;
Best Local Similarity 46.0%; Pred. No. 1.1e-176;
Matches 533; Conservative 186; Mismatches 392; Indels 48; Gaps 13;
QY 11 AFKKIIVANRGEIIVARAPRALETGAATVATYPRREDRGSFHSFASSEAVRITGESSPYKA 70
Db 5 SIQKLVANRGEIIVARAPRALETGAATVATYPRREDRGSFHSFASSEAVRITGESSPYKA 64
QY 71 YLDIDELIGAAKKVAKADAIYPGYGLSENAQIARCAENGITGIGPTPEVLDLDTGDKSRA 130

|||||
Db 65 YLDIEGIIIDIAKKNKVDALHPGVGLSENHFARRCBEEBIVITGKSEHLDMGFGKYKA 124
QY 131 VTKAKAGLPLVLEST-PSKNIDEIYKSAEGQITPIYKAVAGCGRGMRFFVAPDELK 189
Db 125 REQAEKAGIPIVIGSGSPAEETLEAVBOFGANQYPIIKASLGGCGRGMRYSESEVE 184
QY 190 LATEASREAPAGFGDAVYVERAVINPQHEVOILDDHGEVYHLERPOCSIORRQKY 249
Db 185 AYERASSEAKAARAGNDEVYERKLEIENPKHIEVOYIGDKGNVYHLERPOCSIORRQKY 244
QY 250 ETAPAOHLDELDRICADAVKFCRSIGVAGTVEFLVDKGNHVFIEINPIQVEHY 309
Db 245 EVAPSVSLSPELBDQICEAAVALAKUNVYINACTVEFLV-ANNEFYFIEVNPVQVEHT 303
QY 310 TEEVTEVDLYKQOMRLAAGATL--KELGLQOK-ITKGAALOCRTTDEDPNNGFPPDG 366
Db 304 TEMITGVVDIVQIOLVAQGHSLHKKVNIPEODIPTIGYAIQSRYTDEDPODFMPDYG 363
QY 367 TITAYRSPGAGVRLD-GAAQLGSEITAHFMSMLVKMTCRGSDFTAVARAQRALAEFTV 425
Db 364 KIMAYRSGGFGVRLDTGNSFOGAVITPYIDSLVLSWTALTFEQAALKMRYNLOEFTI 423
QY 426 SGVAINIGFLRALLREEDFTSKRIATGFIADPHILOAPPADEGRILIDYADVTVN-- 483
Db 424 RGIKTNIPLENVAKHEKFLTGQYDTSFIDTTPLEFNPKOKDRGKMLTYIGNVTYNGF 483
QY 484 -----KPHGVPRKDVAAPIDKLPIKIDLPPLRGSRDRILKOLGPAARADREDAVAL 536
Db 484 PGIGKKEKPEFDPGLKVDVDOQ-----ARGKQILDEKGAEGLANMWEKESVYL 536
QY 537 TDTTERDAQSLATFVRSEFALKPAEAVAKITPELLSYEANGATDYAMRFLFEDPD 596
Db 537 TDTTERDAQSLATFVRSEFALKPAEAVAKITPELLSYEANGATDYAMRFLFEDPD 596
QY 597 RIDELEAMPNANNIOMLGRNTVGTTPYPSVCARFVKEAASGVDFIRFDALNDVSQ 656
Db 597 RLEDURKEVNPFLFOMLRSSNAVGYTNPVNIKEFYKQASQSIDVFRIDSLNWVG 656
QY 657 MRPAIDAVLETNMAEVMAYSGDLSNENKLYTLDYIKMAEFYVSGAIIILIKMA 716
Db 657 MTLAIDAVDTC-KVAEAAICTGDLIDNRKRYDLATYTSMAKLEDAAGAILIGKDMA 715
QY 717 GLLRPAAYTKLTALRREFDLRVHHTHDYAGQIATYFAAQAQADAVGASAPLSGTT 776
Db 716 GLLRPAAYTKLTALRREFDLRVHHTHDYAGQIATYFAAQAQADAVGASAPLSGTT 776
QY 777 SOPSLSAIYAAFAHRRROTGLSLEAVSDLEPWEAVRGLIYLFESGTFEPTGRVYRHEP 836
Db 776 SOPSLSAIYAAFAHRRROTGLSLEAVSDLEPWEAVRGLIYLFESGTFEPTGRVYRHEP 836
QY 837 GQOLNLRQAATLALGLADREFELIEDNYAAVNMELRPTKVPSSRVYGDIALHVGAGVD 896
Db 836 GQOLNLRQAATLALGLADREFELIEDNYAAVNMELRPTKVPSSRVYGDIALHVGAGVD 896
QY 897 PADPRADPOKIDIPDSVATFLRGLGPNPGGPEPLRTALRGSEKAPLVEPEE-- 953
Db 896 EKDYVEKESIDLPSPSYVELEFGNIGOPHGPEPKILKILKGOE-----PIYVRGELLE 951
QY 954 -----EQAHILADDSKERRNSLRLPLPKPBEFLFENRRRRGNTSALDDREF 1001
Db 952 PYSEFAIOERKEQNLNLEISD-----QDAVAIALYKVTYDTYKTESIGDLSVLDTPPF 1007
QY 1002 YGLVEGRETLIRLPDVRTPLVLRDLAISPPDDKGRNRYVANYNGQIRPMRYRVSSEVT 1061
Db 1008 YGMTLGEIEIWEIERGKI-LIVKLISIEBPQDARRVYVFPENQGPFRVYIKDESIKSSV 1066
QY 1062 ATAERADSSNKGHVAAPRAGVYTVVAC-GDEVKAGDAVAILIEMKMEATTTASVQKID 1120
Db 1067 QERLKADRTNPSHIAASMPGTVIKVLAEGTKVYNGDHLMINEMKMETVQAPFSGTIK 1126
QY 1121 RVVVPAAATKVEGGDLIVYV 1139

Db 1127 QVHKNGEPIOTGDLLEI 1145

RESULT 13

AB47612

ID AB47612 standard; Protein: 1146 AA.

XX

AC AB47612:

XX

DT 05-FEB-2002 (first entry)

XX

DE Listeria monocytogenes protein #316.

XX

KM Antibacterial: gene therapy: vaccine: biosynthesis; biodegradation:

XX

KW vitamin B12; bacterial infection; disease.

XX

OS Listeria monocytogenes.

XX

PN WO200177335-A2.

XX

PD 18-OCT-2001.

XX

PF 11-APR-2001; 2001WO-FR01118.

XX

PR 11-APR-2000; 2000FR-0004629.

XX

PA (INSP) INST PASTEUR.

XX

PI Buchteser C, Frangeul L, Couve E, Rusnlok C, Fsihi H, Deloux P;

PI Dussauget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;

PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JB;

PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;

PI Chakraborty T, Dommann E, Hain T, Berche P, Charbit A, Durant L;

PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;

PI Meduendo E, De Padlos B, Wehlund J, Kaerst U, Entlian K, Hant J;

PI Rose M, Voss H;

DR WPL; 2002-010914/01.

XX

PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment

PT and prevention of Listeria and related bacterial infections, and

PT related polypeptides -

XX

PS Claim 6; SEQ ID No 317; 192pp; French.

XX

CC The present invention relates to the genome sequence of Listeria

CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of

CC it are useful for selecting probes and primers for detecting genes in L.

CC monocytogenes and related organisms, and for studying genetic

CC polymorphisms and other genomes. The present sequence is a protein

CC encoded by the genome sequence of the present invention. Proteins

CC expressed from the genome sequence are useful for raising specific

CC antibodies, identification of L. monocytogenes and related organisms, and

CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin

CC B12. The genome sequence and proteins encoded by it are also useful for

CC selecting compounds that regulate gene expression and cell replication

CC and modulate L. monocytogenes-related diseases. In addition, the genome

CC sequence and proteins encoded by it are useful in pharmaceutical and

CC vaccines compositions for the treatment or prevention of infections by L.

CC monocytogenes and related organisms.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIP0

CC at ftp.wipo.int/pub/published_sequences.

XX

SQ Sequence 1146 AA:

Query Match 43.1%; Score 2493; DB 23; Length 1146;

Best Local Similarity 45.8%; Pred. No. 2,1e-174;

Matches 526; Conservative 189; Mismatches 408; Indels 30; Gaps 11;

OY 13 KKLIVANRGEIIVRAFALEGTGAIVAIIPREDRGSFHRSPASEAVRIGTSGSPKAVYL 72

DB 5 KKLIVANRGEIIVRAFALEGTGAIVAIIPREDRGSFHRSPASEAVRIGTSGSPKAVYL 64

OY 73 DIDEIIGAARKVKADAIYPGVGFLENNAOLARECAENGITFIPPTPEVLDTGDKSRAYT 132

DB 65 DIENITIEIAKESGADALHPGIGFLENIEFARRCEOGKITFVGKSHLDLFGKIKAKE 124

OY 133 AAKRAGLPVLAEST-PSKNIDEIVKSAEGQTPYFVAVAGGGGRGFRVAPSDERKLA 191

DB 125 QALLADIPIVPGSGNPVAGIKVEFEPEKKNYPLMIKASLGSGGRGRVRESKREHVESF 184

OY 192 TEASREAAAFGGAGVYVERAVINPOHIEVOILDHTGEVHLIERDCSLORRQKVEI 251

DB 185 ERASSEAKAAFGNDEVYERKCVNMPKHIEVOILDTGNTVHLIERDCSLORRQKVEI 244

OY 252 AAPAHLPELDRICADAVKRCRSIGOGATVEFLVDENGNHIFIMNPJOVEHYHT 311

DB 245 APCNAITSELNRCICDAAVKLMKNVDYINGTVEFLV-EGDDFYFIEVNPVOVEHTTE 303

OY 312 EYVEVDLVKQMRILAGATLKEGLT---ODKITHGAALQCRITTEDPNNNGFRPDGTI 368

DB 304 MITGIDIVQSLFIADGVALHDOLVAIPKQEDIHHSAGISRTTDEDPNNFDPDGRV 363

OY 369 TAYRSPGAGVRD-GAQLGGEITAHFDSKLVKMTGRGSDFEYAVARAQALAEFVSG 427

DB 364 DYRSTGFGVRLDAGNGFOGTVVTPEYDSLVLKCTWGMTEFOATKRMRLIEFRIRG 423

OY 428 VATNIGFLRLIREDEFTSKRIATGFIADHPHLLQAPPADEQRIIDYADVTVNKPFG 487

DB 424 VKTNIPLLVNRPDRASGNVNTSFDTTPELKEPHIRGTYKTLTIGNVTVNKPFG 483

OY 488 VRPDPVAPIDKLPINIKDLP---LPRGSRDLKQLGPAAFARDLREDAVALVDTTFRD 543

DB 484 IKHND-KPYAERLRKPIRGSOISPTKOILDAKPEGVDVAKQOEVLDTDLTRD 541

OY 544 AHOSLATRVRSKPIFOVADAMAHLLPNMSEFEMKAGATPDVAYRFLNEPWRVLETLRK 601

DB 542 AHOSLATRVRSKPIFOVADAMAHLLPNMSEFEMKAGATPDVAYRFLNEPWRVLETLRK 601

OY 604 AMPNVNIOMLRGNTGTGYPYPSVCRAPFKEAASGVDFRFDALNDVOSMRPIDA 663

DB 602 QIPNVFMOMLRGNAAVGYKNYPRNVIREFKQASGVDFRFDALNDVOSMRPIDA 661

OY 664 VLENTVAEVAAMVSGDLPNEKLYTLDYLLMAEIVKSAHIIAIKDMAGILRPAA 723

DB 662 VREAG-KVEVALCYTGIDDDTFRKTYTIDYKMAKEIVAGNHIIIGIKMAGILLRPA 720

OY 724 VTKLVTLRREEDLPVHVHTHDAGQOLATYFAAQAQADAVGASAPLSCTTSQPSLSA 783

DB 721 AYRLIGELKIDVDVPIHLHTDHTSGNGIYTYAAAVSGVDIVDASSAMSGATSQPSMTG 780

OY 784 IVAFATRRDPTGSLSEAVSDLEPYWEVRGLYLPESGTPGPGRYRHRHIEPGGOLSNL 843

DB 781 LVYGLVGNRQTNLDAQNSQIINHYMEVVRHYVDFDNALNSPOTEYIHEMPEGQYTNL 840

OY 844 RAQATAGLADREFLEIDNVAANVEMLRPKVTPSSKVVGDLALHVGAGVADPADPAAD 903

DB 841 QQQAIAVGLGRMEVEMTYVNVQMGDIKYVTPSSKVVGDLALHVGAGVADPADPAAD 900

OY 904 POKIDIDPSYAFIRGELGNPFGMPPELRLALEGSEKAPLITEVPEEQALHLDADS 963

DB 901 GDTIDPDSYIEFPMGEIGQPYGGFPEKLOLVYKGRGTP---PLTDRGALMEVNVADV 956

OY 964 K-----ERRNSLNRLFPKPTPEEFLEHRRGNMSALDREFEYLVGRETLLI 1012

DB 957 KAELEKMGEPTEKDVISTILYKVLVDYODMINKGDTVLTPTPFYKMGMRIGETIEV 1016

OY 1013 RLDPVTRPLVRLDAISEPDDKGRNVVAVNGQIRPMRVDRSVESTATAERADSSNK 1072

DB 1017 ELEKGLI-LIKLINSIGEPADGRVYIFELNGQPREINQDMVQSTVARRKIDITNP 1075

OY 1073 GHVAAPFAG-VVYTVVAGDEGVKAGDAVAITTEAKMEATITASVDGKIDRVVPAATKE 1131

DB 1076 EHVAGATMGSIYQVAVKGSVKKGDPDLLITAEAKMETITQAPDEGVSSITYSDGPTIE 1135

[illegible][illegible]

Thu Sep 25 08:53:46 2003

us-09-974-973a-19.rag

Page 16

Db 999 TFFNGIRGETLEVOIERGKT-LIIRDEIGEPIIDGNRVLFENLNOGRREVLAKDASIK 1057

Oy 1059 SVATAEAKSDNSCHVAAPFAG-VYVYTAEGSEYKAGDAVAITTEAMKEATTASVDG 1117

Db 1058 SAVOVKOKAEPTNNEOGATMSSSVLOVLVYKRGDVERGODPLITTEAMKETTEAPFAG 1117

Oy 1118 KIDKRVVPAAATKVEGDLIVVY 1139

Db 1118 TVDHIIVEEGEFAISSGDLILEV 1139

Search completed: September 24, 2003, 15:55:26
Job time : 68.5263 secs